DOCUMENT RESUME

ED 271 818 EA 018 587

AUTHOR Shapiro, Phyllis P.; Sheehan, Agnes Teresa
TITLE The Supervision of Student Teachers: A New Diagnostic Tool.
PUB DATE 86
NOTE 17p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Education; Foreign Countries; Higher Education; Preservice Teacher Education; *Rating Scales; *Student Teachers; Student Teacher Supervisors; Teacher Behavior; *Teacher Supervision
IDENTIFIERS Canada; Modified Tuckman Form; *Shapiro Scale; Student Teacher Experience Program (Ontario); University of Toronto (Canada)

ABSTRACT

The student teaching experience can be successful if the student teacher, the host teacher, and the student-teacher supervisor share common goals. Because these goals are frequently not shared, a study was undertaken to develop a scale for the observation of student teachers that would both promote student teacher development and improve the interaction of all three parties relating to appropriate and productive teaching behaviors. Twenty-four student teachers enrolled in the Student Teacher Experience Program at the University of Toronto (Ontario, Canada) completed the Modified Tuckman Form, as did their host teachers and supervisors. The student teachers and host teachers also completed the Shapiro Scale, the scale developed for and tested by this study. To test the validity of the scale's ratings, the 24 student teachers were ability grouped by their supervisors. The scale exhibited strong internal consistency. Student teachers appeared to feel equal needs for help in the four subtest areas measured (creativity, dynamism, warmth and acceptance, and organized demeanor), though host teachers found differences. The scale's validity also appeared strong. The Tuckman Form showed more modest inter-correlations and appeared to assess classroom climate effectively, while the Shapiro Scale assessed specific teacher behavior. (PGD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED271818

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Phyllis P. Shapiro

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Supervision of Student Teachers: A New Diagnostic Tool

Phyllis P. Shapiro, Ed.D.
Consultant in Education

Sr. Agnes Teresa Sheehan, Ed.D.
Professor of Education
University of Toronto

EA 018 587

July - August 1986

ABSTRACT

Shapiro, P., Educational Consultant; Sheehan, Sister T., Faculty of Education, University of Toronto. THE SUPERVISION OF STUDENT TEACHERS: A NEW DIAGNOSTIC TOOL.

Most pre-service programs include both theoretical understandings of and practical approaches to teaching. The student teaching experience is seen as central, for it provides the setting in which theories and methodologies learned can be applied. To be successful, however, the goals of the student teaching experience must be common to the student teacher, the host teacher, and the faculty supervisor. Since this is frequently observed not to be the case, the focus of this study was the development of a scale for the observation of student teachers that would both promote student teacher development and improve the interaction of all three parties with regard to appropriate and productive teaching behaviours. The new scale is described in the context of a study involving twenty-four student teachers and their associated host teachers and faculty supervisors. The Scale's range of application, validity, and reliability are discussed as is the Scale's relationship to an existing parallel instrument. Implications for the improvement of pre-service teacher education programs are outlined.

The Supervision of Student Teachers: A New Diagnostic Tool

Most pre-service programs include both theoretical understandings of and practical approaches to teaching. The student teaching experience is, in this context, generally seen as central, for it provides the setting in which theories and methodologies learned can be applied. To be successful, however, the goals of the student teaching experience must be common to the student teacher, the host teacher, and the faculty supervisor. Since this is frequently observed not to be the case, the focus of this study was the dsevelopment of a scale for the observation and assessment of student teachers that would, in the long run, both promote student teacher development and improve the interaction of all three parties with regard to appropriate and productive teaching behaviours.

In a penetrating critique of twelve different approaches to teacher evaluation Haefele (1980) pointed out the difficulty of establishing reliability of systematic observation, suggesting that "Subjectivity and the propensity for bias remain potential problems" (p. 350). Haefele insists, nevertheless, that evaluation must be based on mutual trust, must be a cooperative effort on the part of teachers and supervisors for the good of students and for the improvement of instruction and must include teacher self-evaluation.

Coker, Medley and Soar (1980) address the question of validity in teacher evaluation with respect to the relationship between mastery of a set of competencies and effective teaching.

> Ideally, evidence should be presented that there is a cause-and-effect relationship between mastery of the competence and effectiveness in the classroom. Evidence that the two are correlated is minimal proof of validity. Even such minimal proof is rarely produced (p. 149).

These researchers thus question the dependability of the common beliefs of experts as a basis for defining the goals of teacher education and setting the criteria for admitting student teachers with the profession.

A survey of evaluation forms used today in faculties of education in order to assess student teacher performance indicates that most of these forms have relatively low levels of validity and reliability and are frequently characterized by both a requirement for highly subjective judgments and a format that is lengthy and ambiguous. In an article by Soar et al (1983), current methods are

found wanting, and the authors advocate evaluation procedures that are performance-based and empirically tested:

> The methods of evaluating teachers that we have used in the past and continue to use today are inadequate. Yet the improvement of education depends on identifying competent teachers, diagnosing incompetence, and measuring changes in competence validly and reliably. (p. 239)

An instrument for evaluating teacher competence was developed and field-tested by Wasserman and Eggert (1980) at Simon Fraser University, British Columbia. Performance-based, it focussed on observable behaviour, related to improvement of pupil learning, reflected educational values, emphasized growth rather than grading, and provided the student teacher with a method of self-evaluation and a gauge of professional growth.

Another competency, field-based student-teacher evaluation program using observation strategies is described by Roberta Woolever (1984). Classroom Observation Guides were developed by Woolever for use in an open space, open concept school. This research involved "the identification of generic teaching competencies and a systematic appraisal of student-teachers with reference to established desirable teaching behaviours" (p. 1). While this approach answers many of the questions raised in the literature regarding the validity-reliability of systematic observation as an evaluative tool, there are serious limitations if the developed guides are applied in more traditional structural classrooms and limited to "more complex classroom teaching behaviours such as management of contracts, learning centres, . . . planning and carrying out a unit of study, arranging a field trip, etc." (p. 2). The long detailed form seems time-consuming and somewhat cumbersome if a large number of student teachers are involved.

In a recent study, Shapiro (1984) found that the Modified Tuckman Form was helpful in identifying strengths and weaknesses of teaching behaviours somewhat indirectly by focusing on teaching styles and classroom climate. The purpose of the present study is the construction of a scale to assess teaching behaviours more directly. As opposed to the lengthy Woolever (1984) Guides, a simpler easier-to-administer rating scale seemed needed. Wasserman and Eggert (1980) as well as Woolever (1985) contended that a "diagnostic rather than a summative evaluative emphasis" (p. 23) should characterize student teaching. This new instrument was designed to provide faculty supervisors and host

teachers as well as the student teachers themselves a clearer and more positive means of diagnosing individual strengths and weaknesses. Much more direct action could then be taken to ameliorate weaknesses and enhance strengths through individual counselling and support.

Procedure

Initially, a new instrument, The Shapiro Scale, was developed. The items of this Scale were developed around the four categories of the Modified Tuckman Form: creativity, dynamism, warmth and acceptance, and organized demeanor. Within each category, ten direct behaviours were specified, each one of which was to be rated by both the student-teacher and the host teacher in terms of the amount of additional help perceived to be required by the student teacher in that particular area of teaching. Each such behaviour area was to be rated along a four-point scale in terms of whether additional help in the area was needed "a great deal", "somewhat", "very little", or "not at all". Examples of the behaviours included in the forty-item Scale are: "preparing lesson plans", "motivating students to learn", "controlling his/her moods" and exercising flexibility in decision making". The entire Shapiro Scale is provided, for the reader's information at the end of this article.[1]

In constructing the Shapiro Scale, the criteria suggested by other investigators were kept in mind. For example, Woolever (1985) recommended that such scales include items based on common sense, expert opinion, and philosophical orientation as well as empirical research and that -- to the extent possible -- desired teaching characteristics should be stated as behavioural competencies. These suggestions were followed by the present authors in developing the Shapiro Scale items although the Shapiro Scale does not sequence items from simple to complex based on anticipated difficulty of mastery (Woolever, 1984). The Woolever Scale, excellent for its purpose, was also designed specifically for an open space, open concept school and is summative in nature. The Shapiro Scale, in contrast, was designed to be used in a more general way --as an initial clinical instrument for the purpose of opening supervisor-teacher discussions about teaching performance. Its focus is primarily formative. In this respect,

[1]The Shapiro Scale is available by writing to the author at 5A-66 Collier Street, Toronto, Ontario, M4W 1L9. Include $2.00 for postage and materials.

the Shapiro Scale focussed more directly on one of the criteria suggested by Wassermann and Eggert (in Haefele, 1980) and that is that the instrument should be capable of use by the student teacher gauging evaluation and growth. Thus, one form of the Shapiro Scale is intended specifically for student teacher self-assessment.

The sample of student teachers involved in this study were all twenty-four enrollees in an experimental program at the University of Toronto's Faculty of Education. The program, STEP -- Student Teacher Experience Program -- provides students with the opportunity to (a) acquire an understanding of children and how they learn, and (b) understand, practice, and experiment with a range of teaching strategies and materials. A full description of the Program can be found in the special booklet prepared by the sponsoring Faculty (cf. Elementary Education Department, 1984-1985). For present purposes, it need only be added that within this program, the students are placed at two elementary schools at which they observe, do their practice teaching as an ongoing activity throughout the year, and take university courses given by university faculty who, however, come to the schools for the purpose of this instruction.

In the first semester of the student teachers' placement in the program, the faculty supervisor, host teacher, and student teacher completed the Modified Tuckman Form, each describing the student teacher using the bipolar adjectives which comprise the twenty-eight semantic differential items of the Tuckman Form. Both host teachers and student teachers also completed the Shapiro Scale. Finally, the university faculty member who was responsible for running the STEP Program and who was familiar with each of the twenty-four students in terms of their professional work within the program placed each of the students in one of four groups, groups which differed in terms of the ability level of the student in his/her teaching practice. That is, the head of the STEP Program divided the twenty-four student teachers into four groups of six, group one representing, in his view, the best student teachers; group two the second best, etc. etc. This step was taken to develop an initial index of the validity of the Shapiro Scale in terms of the relationship between the Shapiro Scale ratings and an independent assessment of the student teachers' performance.

Results

The data collected in the study were analyzed with a view to addressing the following issues: (a) the characteristics of the Shapiro Scale in terms of its

internal consistency, the relationship among its subtest scores, and the area of greatest need as identified by both the host and student teachers; (b) the validity of the Shapiro Scale in terms of its relationship with the assessment of the host teachers (and the student teachers themselves) and an independent assessment of the student teacher; (c) the factor and subscale structure of the Modified Tuckman Form; and (d) the relationship between the Modified Tuckman Form and the Shapiro Scale. The results of these analyses are reported below.

Shapiro Scale Characteristics

The internal consistency of the Shapiro Scale was assessed through the calculation of the Hoyt estimate of reliability for both the subtest and total test scores. The Hoyt estimates suggested very strong internal consistency for the scale in both respondent groups. The Hoyt estimates for the total test score were .93 and .98 for the host and student teachers respectively. The Hoyt reliability estimates for the Shapiro subtest scores were also high -- generally in the 80s for the host teachers and uniformly in the 90s for the student teachers.

The intercorrelations between the subtest and total test scores on the Shapiro Scale -- for the host and student teachers -- are given in Table 1. This table reveals that the relationships in question are very high for student teachers (the lowest is +.80) but more moderate for the host teachers (probably a more discriminating group) for whom the correlations range from +.60 to +.92.

Table 1 about here

These data indicate that from the student teachers' point of view the same amount of help is required on all four subtest areas of creativity, dynamism, warmth and acceptance, and organized demeanor. The host teachers, however, appear to perceive the individual student teachers' areas of need more differentially, discriminating as between the four aspects of teaching behaviour, thus suggesting that more help is required in some areas more than others. This might be explained, of course, by the host teachers' years of experience in the teaching field.

On the Shapiro Scale, the higher the individual's score, the greater is the perceived need for help in that teaching area. The maximum scale score (i.e., the greatest possible perceived need for help) for any subtest is forty. The total

score is simply the sum of the subtest scores, and the maximum total test score, is therefore, one hundred and sixty. The minimum scores are ten and forty for the subtests and total test scores respectively. The Shapiro Scale means and standard deviations are presented -- for both student and host teachers in Table 2.

Table 2 about here

These data suggest a considerable variability on each of the subtest scales with regard to the perceived need for assistance. In general, the perceived need was higher from the point of view of the student teacher than from the perspective of the host teacher, but this tendency was marked only in the areas of organized demeanor and creativity. The differences on dynamism and warmth and acceptance were very small although even in these areas, the variability is greater in student ratings than in teacher ratings.

Validity of Shapiro Scale

In order to develop an initial sense of the validity of the Shapiro Scale scores, these scores were compared (as outlined above) to the assessments of the same student teachers by an independent evaluator. The independent evaluator, the STEP program director, had worked with each of the 24 student teachers in a methods of teaching class and, in addition, had observed their classroom practice. On the basis of this experience, he placed the student teachers into one of four groups depending upon his global perception of the students' teaching skill. The means of each of the Shapiro scales for each of the independent evaluator's groupings are presented in Table 3. It is evident from the data in this table that the higher ranking groups (i.e., the better student teachers) are seen -- by both the host teachers and the student teachers themselves -- as those least in need of help.

Table 3 about here

This general pattern of agreement between the independent evaluator and the Shapiro Scale's indices was at least partially supported by the results of a Newman-Keuls analysis (a test of multiple comparisons between means) which yielded statistically significant results (at the five percent level) when the most

skilled student teachers (i.e., Group One) were compared with the least skilled (i.e., Group Four).

Structure of the Modified Tuckman Form

In terms of the data from the Modified Tuckman Form the item inter-correlation matrix provided data suggesting the integrity of the factor structure as originally described by Tuckman (1976). That is, the items on the various sub-scales (e.g., creativity, dynamism, etc.) cohered appropriately and seemed relatively independent of each other. Thus, the inter-correlations between the Tuckman scales (with the exception of organized demeanor and warmth and acceptance) are more modest than those for the Shapiro scales. As illustrated in Table 4, these correlation coefficients again provide some confirmation of the relative independence of the Tuckman dimensions.

Table 4 about here

Further examination of the Tuckman data in terms of their relationship to the evaluator's grouping of the student teachers did not produce for either the faculty supervisors, host teachers, or student teachers any statistically significant relationship between the Tuckman measures of classroom climate and the independently assessed teaching skill of the student teacher. To some extent, this may be due to the fact that the student teacher is operating in a classroom whose climate has been primarily developed by the host teacher. Alternatively, the data may simply underline the indirectness of the approach of the Modified Tuckman Form as compared to the Shapiro Scale.

The Tuckman Form did, however, yield consistent data of another sort. All student teachers whether as rated by their faculty supervisor, the host teacher, or themselves appeared to rank higher on organized demeanor and warmth and acceptance than on creativity or dynamism. This result might, of course, be expected since pre-service education programs orienting students to the practice teaching experience often focus on both the importance of positive reinforcement (i.e., warmth and acceptance) and careful preparation (i.e., organized demeanor). Dynamism and creativity are more likely to emerge when basic teaching behaviours are consolidated. The Woolever (1984) graduated model seems to address this general pattern in student teacher behaviour.

Relationship Between the Tuckman Form and the Shapiro Scale

For both the host teacher and the student teachers, the correlation between the parallel Shapiro and Tuckman subscales were not high. The range for the host teachers was +.33 to +.42 while for the student teachers, it was +.36 to +.51. This suggests, of course, that the Tuckman and the Shapiro Scales are not appropriate substitutes for one another. The Modified Tuckman Form assesses classroom climate whereas the Shapiro Scale addresses specific teacher behaviour more directly.

Discussion

In general, the host teachers responded favourable to the Shapiro Scale primarily because, unlike the Tuckman Form, it focussed on specific teaching behaviours to which they could more directly relate the observed behaviour of the student teachers. Typical comments were "It was easy to use", "It was not wordy", "It was very clear", etc., etc. It was not, therefore, entirely surprising that the results from the Shapiro scales so closely corresponded to the independent assessment of the teacher practice of the student teachers.

The Shapiro Scale is, therefore, clearly of some value in the evaluation of student teachers. It is geared more to growth than to grades and for this reason is non-threatening to student teachers. Woolever (1984) on the other hand considers her Guides to be "invaluable as a record of student teacher progress or lack thereof. For example, they can be used as evidence to document that a particular student . . . should not receive a passing grade." Both approaches do include the collaboration of students, host teachers and faculty supervisors to the benefit of the student teacher.

More work is, however, needed to improve the scale. If, for example, the Shapiro Scale is to be useful in the self-assessment of student teachers, further item development will be required to (i) yield subscales more independent of each other and (ii) replace some items with behaviours more likely to be relevant to the special case of the student teaching environment. Further, relative to the reliability of the Shapiro Scale, internal consistency estimates are not fully adequate. Inter-rater reliability must also be established. Relative to future developments, a standard replication of the present study would be useful as would some attempt to assess the changes in student teacher behaviour over time and in alternative program contexts.

The development of the Shapiro Scale is, of course, not an end in itself. It can only serve as a first step in focussing on student teachers' strengths and weaknesses. A remediation program would ensue which would help alleviate weaknesses and build on strengths. The Scale could then be reapplied to assess the degree to which the desired changes actually occurred. Finally, the Shapiro Scale might also be very helpful for in-service development of practicing teachers, but that is an area which would require an independent investigation and inquiry.

References

Coker, H., Medley, D.M. & Soar, R.S. (1980). How valid are expert opinions about effective teaching? Phi Delta Kappan, 62, 131-134, 149.

Elementary Education Dept. F.E.U.T. (1984-85). Primary Junior Handbook. Roden/Corpus Christi Project.

Haefele, D.L. (1980). How to evaluate the teacher - let me count the ways. Phi Delta Kappan, 61, 349-352.

Shapiro, Phyllis P., Cohen, Phyllis, Wright, Nancy & Pollock, Gloria. (1984). Supervision: A new approach to in-service teacher education. Education Canada, Fall, 22-25.

Soar, Robert, Medley, Donald & Coker, Homer. (1983). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan, 65, 239-246.

Tuckman, B.W. (1976). New Instrument: The Tuckman teacher feedback form. Journal of Educational Measurement, 13, 233-237.

Woolever, R.M. (1984). Observing student teachers for a hierarchy of generic teaching skills. School of Education, University of North Carolina at Chapel Hill. (ERIC Document Reproduction Service No. (238 839)

Woolever, R.M. (1985). State-mandated performance evaluation of beginning teachers: Implications for teacher educators. Journal of Teacher Education, 36 (2), 22-25.

Table 1

Shapiro Scale: Correlations Between Subtest and Total Test Scores

Scale	Host Teachers (N=24)				Student Teachers (N=24)			
	Org	Dyn	W/A	Creat	Org	Dyn	W/A	Creat
Dynamism	.60				.84			
Warmth & Acceptance	.75	.80			.88	.85		
Creativity	.73	.57	.69		.80	.81	.82	
TOTAL	.87	.86	.92	.85	.93	.94	.95	.91

Table 2

Shapiro Scale: Means and Standard Deviations

Scale	Host Teachers (N=24)	Student Teachers (N=24)
Organization	21.25*	26.63
	(4.99)**	(6.18)
Dynamism	21.79	21.92
	(6.23)	(7.71)
Warmth & Acceptance	19.29	21.96
	(5.41)	(6.61)
Creativity	22.04	26.54
	(5.87)	(6.30)
TOTAL	84.38***	97.04
	(19.71)	(25.13)

*Maximum Score of subtests (i.e., in need of most help) is 40

**Figures in parentheses are the standard deviations

***Maximum Score of TOTAL (i.e., in need of most help) is 160

Table 3

Mean Shapiro Scale Scores for Evaluator's Grouping of Student Teachers

Shapiro Scales: Host Teacher Ratings

Evaluator Grouping (student teaching skill)	N	Creat	Dyn	Org	Warm/Accept	Total
1 (highest)	6	21.1	17.8	17.8	16.5	73.5
2	6	19.3	21.8	20.1	17.8	79.1
3	6	21.8	21.8	23.3	19.0	85.3
4 (lowest)	6	25.8	21.8	23.6	23.8	99.5

Shapiro Scales: Student Teaching Ratings (Self-Assessment)

1 (highest)	6	21.6	17.3	22.3	18.5	79.6
2	6	26.0	22.3	26.6	19.5	94.5
3	6	29.3	21.6	28.3	24.0	103.5
4 (lowest)	6	29.1	26.3	29.1	25.8	110.5

Table 4

Inter-Correlation Matrix of the
Tuckman Subscales

	Dynamism	Organiz	Warmth/Accept
Faculty Supervisor (N=24)			
Creativity	.58	.64	.59
Dynamism		.14	-.01
Organization			.92
Host Teachers (N=24)			
Creativity	.41	.48	.20
Dynamism		.44	-.07
Organization			.60
Student Teachers (N=24)			
Creativity	.42	.27	.29
Dynamism		.62	.47
Organization			.75